Technical Note

1968-8

A Study of Two Short-Period Discriminants

E. J. Kelly

12 February 1968

Prepared for the Advanced Research Projects Agency
under Electronic Systems Division Contract AF 19(628)-5167 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Lexington, Massachusetts

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This research is a part of Project Vela Uniform, which is sponsored by the U.S. Advanced Research Projects Agency of the Department of Defense; it is supported by ARPA under Air Force Contract AF 19(628)-5167 (ARPA Order 512).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

A STUDY OF TWO SHORT-PERIOD DISCRIMINANTS

E. J. KELLY

Group 64

TECHNICAL NOTE 1968-8

12 FEBRUARY 1968

LEXINGTON MASSACHUSETTS

ABSTRACT

This paper is the result of a large-population study of discrimination between earthquakes and explosions, using only short-period data from a single station. The data was obtained from the Large Aperture Seismic Array and the two discriminants studied were waveform complexity and spectral ratio. Procedures for multivariate discrimination are developed and results are given in terms of earthquake identification percentages using these discriminants singly, and in combination with each other and with magnitude. The results are quite encouraging, complete separation being found for explosions and shallow earthquakes using spectral ratio and magnitude together.

Accepted for the Air Force
Franklin C. Hudson
Chief, Lincoln Laboratory Office

I. INTRODUCTION

The problem of discriminating between earthquakes and explosions by seismic means has progressed to the point where positive separation is clearly possible for large events. Barring unforeseen anomalies in the behavior of explosive sources, detonated in untried, aseismic regions, this separation can be provided by comparing the relative excitation of long-period surface waves to that of short-period P-waves[1,2] and the threshold of effective discrimination by this means appears at present to be at a body-wave magnitude somewhere in the range 4 3/4 to 5. These results are based upon data from world-wide networks of stations, but very similar results have been reported for Central and Eastern Asian events recorded at a single station, the large aperture Montana LASA[3] array, by Capon, et al,[4] who quote a conservative discrimination threshold of 4.9 (body-wave magnitude). The problem that remains is to lower the threshold of discrimination to a value closer to the threshold of detection of P-waves for a given network or array station. The surface wave – P wave method is presently limited by the difficulty of detecting the long-period waves from weak events, and many people are now concentrating their efforts on improving this detectability.

Recently, however, it has been found by Briscoe[5] that very good separation can be obtained at a single station from short-period data alone, by plotting the spectral ratio versus P-wave magnitude (spectral ratio is the ratio of received P-wave energy in the 1.4 to 1.9 Hz band to that in the 0.4 – 0.8 Hz band). This has caused a revival of interest in short-period discriminants, and the present investigation was undertaken

in order to collect a population of Sino-Soviet earthquakes, using LASA data, and to study spectral ratio, as well as the older discriminant, complexity, with three objectives:

(1) To obtain single-station performance estimates of these discriminants, used singly and in combination, on a large population of Sino-Soviet events,

(2) To determine the magnitude threshold at which effective discrimination can be achieved using single-station short-period data alone, and

(3) To learn what, if any, advantages there are in the use of a large short-period array for discrimination purposes, other than its obvious ability to enhance signal-to-noise ratio.

A summary of our conclusions, not exactly unequivocal, will be found in Section VI.

II. THE DATA BASE

In order to collect data from events in the Soviet Union and China, arrangements were made with the Data Center in Billings to send us routinely recordings of all events for which epicenters in this region were obtained by the usual triangulation procedures. After about two months, only recordings of events showing no evidence of the existence of pP on the raw data were sent. In this way a library of over 200 events was formed and further processing initiated. Data continues to be collected, of course, for future extension of the study.

The recordings forming the raw data base were subjected to a sequence of measurements, beginning with a redetermination of epicenter by beamsplitting. An initial pattern of beams was formed from the 21 subarray straight sums. The pattern had a spacing of about 2^{o} and was centered on the preliminary epicenter reported in the LASA bulletin (sometimes a starting epicenter was redetermined by triangulation). The coordinates of the best beam were taken as a revised epicenter. If this epicenter was near the edge of the pattern, a new pattern was formed centered on the revised epicenter and, in most cases, a finer pattern was formed (with 1^{o} and sometimes 0.5^{o} spacing) to obtain a final result. Events too weak to permit epicenter determination in this way were removed from the data base. Events were retained if their epicenters were within a few degrees of the Soviet Union, China or Eastern Europe, including Greece and Turkey. About 100 earthquakes remained at this point. Figure 1 shows a rough map of this area, together with the earthquake and explosion epicenters.

Once an epicenter was determined, a new beam was formed, as follows. Steered subarray beams were first formed, using only the 4, 7, and 8 rings, a total of nine elements, from each subarray. These subarray beams were then combined, with appropriate steering delays and the best available station corrections, to form the final beam. Dead or excessively noisy channels were automatically deleted from the beam. The output of this beam, both raw and bandpass filtered (with a 0.6 to 2.0 Hz pass band) was played out on a chart recorder, along with the individual steered subarray beams, each delayed by the amount used in forming the beam. If the revised epicenter was still in error, or if the station corrections used were inadequate, then the failure of subarray alignment would be observed and corrected, using hand-picked corrections to guarantee alignment.

As a result of the processing up to this point, playouts were available of the filtered and unfiltered beam waveforms, representing some 15 to 18 db of signal-to-noise ratio gain in the signal band, relative to the individual raw recorded seismograms.[6] P-phase amplitudes and dominant periods were manually read from these waveforms, permitting a measurement of body-wave magnitude, and a search was made for pP. Any secondary P-phase within two minutes of P and not associated with the expected arrival time of PcP was considered a possible pP. If the amplitude of such a phase exceeded either that of P itself or 1.5 mμ (on the normalized beam), and if the phase delay after P exceeded four seconds, then the phase was called pP and the event was considered "deep." Events not showing this evidence of depth were retained as

"shallow." We realize, of course, that single-station pP determination is risky, and that in practice one would rely heavily on the P-arrival times at a world-wide network of stations. However, only about half of the earthquakes in our population were identified by the U. S. Coast and Geodetic Survey (USCGS), and many of these were located at the non-committal 33 km depth (USCGS depths, when reported, agreed well with our determinations). Hence, we relied on our own measurements in order to screen out "deep" events as earthquakes. About half the initial population was considered deep by this test, but the subsequent discrimination analyses were done on both the screened ("shallow") population and the original total population for comparison.

In Figure 2 we show the magnitude distribution for the initial population of 93 earthquakes for which a final beam was formed. The data is presented here in the form of a cumulative seismicity curve, in which the logarithm of the number of events in the population having a magnitude greater than some value is plotted against that magnitude value. Thus the curve rises for smaller magnitudes until the logarithm of the total number of events is reached. In all the remaining histograms in this report, we plot the percentage of the population, on a linear scale, having a parameter value less than or equal to a given number versus that number. Figure 3 is such a histogram of epicentral distance for the same events used in Fig. 2. Events more than 95^{o} from LASA have been removed from the population and not subjected to further processing, in order not to prejudice the study by data beyond the normal range at which a station would be expected to provide useful information.

We have tried not to screen our population in any way which would affect its suitability as a typical sample for the discrimination analysis. The magnitude distribution looks quite reasonable, and the leveling off at small magnitude in Fig. 2 is indicative of our threshold for detection and location. Figure 3 describes a broad and fairly uniform distribution of distances, although natural seismicity has concentrated much of the data in the Kurile-Kamchatka region. The depth screening done at the site during part of the experiment may unbalance the population slightly, but we found phases identified as probable pP on beam outputs at about the same rate after as before the beginning of this screening period.

A list of the earthquakes used in this study may be found in the Appendix.

Nineteen events of possible or confirmed explosive origin (hereafter referred to collectively as "explosions") were used in the study, five of which occurred during the data collection period itself. The others are older recordings, altogether representing four source points: (1) Amchitka, the epicenter of the Longshot event; (2) Novaya Zemlya, the source of a large Soviet shot; (3) Semipalatinsk, the source of a large number of possible Soviet explosions; and (4) Southern Algeria, the site of two possible French explosions.

A serious deficiency in our present population of events is the disparity in the magnitudes of the earthquakes and the explosions. All of the explosions exceed magnitude 4.9, while the initial population of earthquakes had an average magnitude of 4.7. Magnitude histograms of these two populations are given in a composite plot in Fig. 4.

In order not to present only results which might be seriously biased by the magnitude differences between earthquakes and explosions, we have used both the total earthquake population and the smaller number of earthquakes having magnitudes not less than 4.8. We previously mentioned that we were using both the total population and the apparently shallow events, hence the discrimination performance results given below in Section V are quoted for four earthquake populations as shown in Table I.

TABLE I

All earthquakes*	85 events
Large earthquakes**	32 events
Shallow earthquakes+	48 events
Large, shallow earthquakes	12 events

---

* All events, at distances not greater than 95^{o}, for which a final beam was formed and all discriminant measurements were made.

** Measured magnitude equal to or greater than 4.8.

+ pP phase not identified on beam output.

III. THE DISCRIMINANTS

For the purposes of this study, the short-period discriminants considered were depth (pP detection), waveform complexity, and spectral ratio. The depth determination and pP detection criterion were described in Section II. The remaining two are described in detail here. Of course, location is a primary discriminant, but this study began with a population of events which were just from the area that would presumably be monitored (Eastern Europe, the U.S.S.R. and China). Moreover, location is best determined by a network of stations, and one of our purposes was to assess the contribution of a single array station to the discrimination operation by means of its ability to measure waveform discriminants on processed (i.e., high signal-to-noise ratio) traces.

It has often been observed that explosion seismograms, at teleseismic distances, are impulsive and relatively simple. Nearly all the energy that is ultimately received arrives in the first several seconds. While some earthquakes share this property, a great many are more complex, with P phases lasting from many seconds to minutes. Moreover, the simplest earthquake waveforms are usually associated with very deep events which would easily be removed from consideration at an early stage of the screening process. In order to get a quantitative measure of waveform complexity, the British seismic discrimination group (at UKAWRE) has suggested[7] a particular functional of the waveform, namely the inverse ratio of the energy received in the first five seconds of the P phase to that received in the subsequent 30 seconds. The durations

of the two time intervals represent arbitrary, but reasonable, choices. However, the results can be expected to depend upon several details such as (1) the processing which went into the formation of the trace in question, (2) the procedure used to determine the start of the first interval, and (3) the manner in which the energy is defined. Moreover, the British definition was always applied to the cross-correlation function between two steered beams, formed from different portions of the total array. We have adopted what we consider to be the essential feature of their definition so as to apply it to a single beam output. In this study the complexity measurements were made on the band-pass filtered (0.6 to 2.0 Hz) beam waveform, and the starting point for computer processing was determined manually (using a light pen on a scope display).

The energy was defined in two ways, as the integral of the square of the signal (quadratic complexity) and as the integral of the absolute value of the signal (linear complexity). The use of the less physical linear measure was motivated originally by dynamic range problems in a fixed-point computer. Taking a time origin at the start of processing, and letting x(t) represent the seismic waveform as a function of time in seconds, then the two definitions are as follows:

Quadratic Complexity

$$C_Q = \int_5^{35} [x(t)]^2 \, dt \Big/ \int_0^5 [x(t)]^2 \, dt \ ,$$

Linear Complexity

$$C_L = \int_5^{35} |x(t)| \, dt \Big/ \int_0^5 |x(t)| \, dt \ .$$

In practice, of course, the "integrals" are actually sums over the data samples, in our case, 20 samples per second. If x(t) were a constant, either definition would yield a complexity of 6, the ratio of interval lengths. We may expect that for pure seismic noise, values like 6 will be observed. We find that for complex and emergent earthquakes, much larger values are attained, while simple events attain values less than unity on both definitions. Figure 5 is an example of a filtered beam seismogram showing the complexity measurement intervals and results. Histograms of the two complexities, as measured on 85 earthquakes, are shown in Figs. 6 and 7.

We also expect an intimate relation between the two definitions. On dimensional considerations alone, we may expect a relation of the form

$$\frac{1}{T}\int_0^T [x(t)]^2\, dt = K\left\{\frac{1}{T}\int_0^T |x(t)|\, dt\right\}^2$$

to hold in some average sense. The above relation is true literally for a constant x(t), and true as a relation between expectation values for a random noise model of x(t). Since the constant K drops out in the ratio of complexities, we can expect, roughly, a relation of the form

$$C_Q = (C_L)^2/6 \ .$$

In Fig. 8 we give a scatter diagram of C_Q vs C_L measurements for the same 85 earthquakes used in Figs. 6 and 7, with the simple relation derived above plotted as the solid curve. It appears that the relation between the two complexities is sufficiently

deterministic to permit us to dispense with one of them, and we have chosen to retain the linear definition, mainly to facilitate comparison with earlier measurements made on LASA data. The two definitions have also been compared in terms of their discrimination capability with the same result, that they are equally effective.

In Fig. 9 we present a composite plot of four histograms of the linear complexity. One curve is for the total earthquake population, another for the earthquake population with the deep events removed and a third in which the Kurile-Kamchatka earthquakes are also removed. The fourth curve is for the explosion data. The fact that shallow earthquakes yield a smaller value of complexity is due to the fact that for most of the deep events, the phase identified as pP falls inside the second time interval used in the measurement, hence increasing the measured complexity value. Events deep enough so that pP does not interfere with the measurement usually have low complexities. The further shift towards small complexity values for continental earthquakes may be significant, but the population involved (shallow, non-Kurile-Kamchatka events) is small and heterogeneous, and may not represent a good sample of continental Sino-Soviet events (part of that area is beyond 95^{o} from LASA).

The curve showing the smallest average complexity is the histogram for our population of 19 explosions. The average value is very close to unity, and actually only four explosions exceeded that value. The really anomalous event was the Novaya Zemlya shot of 27 October 1966 with a measured complexity of 2.7. Other studies of this event show that the relatively long P coda consists of energy arriving from the

same location as the source. The next largest value was attained by Longshot, at 1.2. The separation of these populations for discrimination purposes is discussed in Section V.

Background studies of the spectral character of seismic signals and the motivation for the particular choice of definition of spectral ratio used here are given in Reference 5. The measurement procedure used in this study begins with the unfiltered beam waveform, which is passed through a parallel bank of fifty 0.1 Hz wide digital bandpass filters. These filters have a [sin x/x] -type frequency response, and the absolute values of their outputs are smoothed and sampled once per second and used to form a sonogram on a scope display. This is a display which shows the variation of spectral density with time. Frequency is the ordinate, time is the abscissa and spectral density (the filter output) is represented by the intensity of the display. An example is shown in Fig. 10. Using a light pen, an operator chooses a start time as close as possible to the onset of P, and a stop time defining an interval of 30 – 35 seconds avoiding, if possible, pP or other phases or data errors. The filter output samples falling inside the processing interval chosen by the operator are summed to provide a spectrum of the chosen portion of the waveform. The five spectral values corresponding to the frequencies 1.5 Hz through 1.9 Hz are then summed to form the numerator, and the values corresponding to the frequencies 0.4 Hz through 0.8 Hz are summed to form the denominator of the output spectral ratio. In Fig. 10 the start and stop times are shown by cursor lines, and the spectrum obtained from this time base is shown as the dotted curve above the sonogram.

Separate studies have shown that the measured value of spectral ratio is not critically sensitive to the time interval used in the measurement, but the intervals used were standardized as much as possible. In one relatively bad test case, the spectral ratio decreased by 10% as the length of the measurement interval increased from 25 to 45 seconds.

Measurements of the spectral ratio of pure seismic noise gave an average value of 0.158 (with a range from 0.057 to 0.317), while earthquakes yield a broad distribution, as shown in the histograms of Fig. 11.

The four populations used in Fig. 11 are the same as were used in Fig. 9. This time the explosions have high values, relative to the earthquakes, and a more distinct regional dependence for earthquakes is apparent. It is interesting that the removal of the deep events from the total earthquake population simply removes the tail of the distribution at high values of spectral ratio, without changing the shape of the remaining part. The details of the population separation by means of spectral ratio are given in Section V.

A much less sensitive characterization of the spectral content of an event waveform is provided by the dominant period. We measured these periods, in the traditional way (time between peaks), on the bandpass-filtered beam waveforms and compared the results with the corresponding measurements of spectral ratio. Figure 12 is a scatter diagram of dominant period versus spectral ratio for our full earthquake and explosion populations. The correlation is rather poor, except that the value of

period seems to set an upper bound to the spectral ratio. Very similar results are obtained when only the shallow earthquakes are used. We have also noticed a poor correlation between dominant period and the peak frequency of the corresponding spectral density, although numerical data are not now available.

Two spectra, one from an earthquake, one from an explosion are shown in more detail in Fig. 13. The similarity of the spectra at frequencies above the peak frequency is typical of the two populations. The main difference seems to be the sharp drop in signal energy at the low end for explosions, compared to earthquakes. It is perhaps useful to think of the spectral ratio as an inverse measure of low frequency content, normalized by the relatively stable high frequency portion. In a similar way, the complexity is a measure of the highly variable coda content, normalized by the primary energy in the initial part of the P phase.

It is clear that reliable measurements of complexity and spectral ratio require a certain minimum signal-to-noise ratio. The complexity is likely to be too high for weak, simple events because the coda is weaker than the noise, while the spectral ratio is affected by the non-uniform noise spectrum, especially by the usual low frequency peak. It is difficult to give accurate values to these limits, in magnitude units, but for the earthquakes studied in this work it appears that complexity and spectral ratio are being reliably measured on the signal component of the seismogram only for magnitudes greater than 4.3 for our present methods of processing. Reliable measurements are obtained for all the explosions in our data base because they are all large.

One of the most interesting questions to be resolved when we obtain data from weaker explosions is the analogous magnitude threshold for reliable measurement of the short period discriminants.

IV. DISCRIMINATION PROCEDURES

There exists a considerable literature devoted to the study of statistical techniques for multivariate discrimination between two or more populations. However, a common feature of nearly all this work is a statistical model governing the behavior of the variables used for discrimination under different hypotheses. In our case we have no basis for making such statistical assumptions and the measurements made so far (on rather small populations) do not suggest any simple joint probability densities among the discriminants used here. Instead we have adopted a simple principle of exclusive regions, which is best viewed geometrically in the following way.

Suppose that a set of M measurements (such as body-wave magnitude, surface wave magnitude, spectral ratio, etc.) is made on each event of two large populations, one consisting of earthquakes and the other consisting of explosions. We can think of each set of measurements as a point in an M-dimensional space, so that the earthquakes constitute one swarm of points and the explosions another. In many similar problems such swarms for different populations would overlap significantly so that separation can only be accomplished at the cost of occasionally mistaking each population for the other. However, it usually turns out when explosions and earthquakes are compared that there is a limited region of population overlap, a sizeable region containing only earthquakes and, in some cases, a region containing only explosions. In general, the earthquakes represent a large family of waveforms compared to explosions, and, with respect to some parameters, the explosion population is completely contained within that of the earthquakes.

Based on this observation, our procedure is to define a region, R_X, in the M-space which is just large enough to contain all the explosion points, and another, R_Q, which is just large enough to contain all the earthquakes. These regions can be defined in various ways, one convenient definition being the smallest convex body bounded by hyperplanar surfaces which includes the points. We now divide the entire space into four regions, as follows:

$$R_1 \equiv R_X - R_Q \qquad \text{(points in } R_X \text{ but not in } R_Q\text{)},$$

$$R_2 \equiv R_Q - R_X \qquad \text{(points in } R_Q \text{ but not in } R_X\text{)},$$

$$R_3 \equiv R_X R_Q \qquad \text{(points in both } R_X \text{ and } R_Q\text{)},$$

$$R_4 \equiv M - R_3 \qquad \text{(points in neither } R_X \text{ nor } R_Q\text{)}.$$

We envision using these regions to classify a new measurement by saying that if the corresponding point falls in R_1 that "explosions are known to look like that, but no earthquake in our experience does," and a converse statement if the point falls in R_2. In the case of R_3 or R_4 we are non-committal, in the former case because such measurements have been observed for both types, and in the latter because we have no observations in that part of the space at all. If a measurement falls in region R_1 or R_2, it is tempting to make a quantitative statement of confidence in the decision made. However, in the absence of assigned probability measures, any numerical assignment of confidence would be highly arbitrary. Even so, if one makes a measurement and the corresponding

point lies in the earthquake-only region (R_2) and is very far from the nearest boundary of R_x (i.e., far from any explosion point), one feels more confident of his decision than if the measurement lies near R_x. We shall not pursue the point further in this report.

The procedure described above in somewhat mathematical terms should, of course, be used only as a guide, but some sort of rough separation with exclusive regions should be attempted. In particular, we are interested in regions R_1 and R_2, since an event falling in either is then positively identified. Of course, the entire separation procedure should be preceded by such binary screening operations as tests for epicenter (in or out of the region being monitored) and depth (any reliable seismic evidence of depth puts the event out of the explosion category).

This procedure may be used as a way of combining several measurements, each of which can meaningfully be used separately as a single-variable discriminant. However, in the multi-dimensional case, measurements can be included which have no discrimination significance by themselves, such as magnitude. In this case it is the nature of the dependence of one or more discriminants on magnitude which is diagnostic of the source.

In the one-dimensional case, i.e., discrimination based on the value of a single quantity, the regions described above reduce to simple intervals. Thus, taking complexity as an example, both explosions and earthquakes attain values near zero, but the largest value for explosions is 2.7. Hence, to use this measurement alone, we would

classify an event with complexity greater than 2.7 as an earthquake. To be more conservative, one would probably increase this threshold a little, and be non-committal regarding any value greater than 10.0 (region R_4). In this case there is no region R_1 because explosions attain no values not also attained by earthquakes.

Having established these regions for one pair of reference populations, one would like to test them on a second pair. Given only the reference population itself, we can only establish the four regions and see how many earthquakes fall in R_2 and how many explosions fall in R_1. The corresponding percentages can then be interpreted as estimated performance characteristics. In other words we can only say, on the basis of regions established on one pair of populations, that a certain percentage of the earthquakes and another percentage of the explosions of a second very similar pair of populations would be correctly identified. It is in this sense that the discriminants are evaluated, singly and in combination, in the next section.

It would be desirable to have a definition of what constitutes satisfactory discrimination performance. Since perfect separation is clearly desirable, one is really asking how useful imperfect separation is, and this in turn depends upon the political ground rules. If one fails to identify an explosion, then the entire system has simply failed to function. If this happens only rarely, then such an error may be tolerable, since one can still gauge the level of testing activity of the countries being monitored. However, in most discussions it is assumed that this error is unacceptable, and that all events are explosions until proven to be earthquakes. The error of failure

to identify an earthquake as such may now lead to an inspection and the permissible rate of falsely triggered inspections, while no doubt very small, is totally dependent on the terms of a treaty. It is, therefore, beyond the scope of this report to give a numerical definition of satisfactory discriminant performance and we shall let the reader judge the results for himself. It must be kept in mind that our performance percentages refer to various discriminant combinations as applied at a single station. We do not suggest that discrimination be performed, in actual operation, using data from one station if it is at all possible to obtain measurements at other stations of a network.

V. DISCRIMINATION PERFORMANCE

We have already described, as an example in Section IV, how complexity would be used alone as a discriminant. A threshold complexity is chosen, slightly higher than that of the most complex explosion, and events having larger values are classed as earthquakes, i.e., as being in the R_2 region. There is no classification of events as explosions (no R_1 region), since there is no evidence that explosions attain values lower than those ever attained by earthquakes. Identification percentages, based on our populations themselves, are given in Table II.

The use of spectral ratio alone as a discrimination procedure is closely analogous, except that now high values are associated with explosions. A threshold near 0.425 (the lowest spectral ratio observed for explosions) is chosen to form the boundary of the R_2 region, and events with smaller values are classed as earthquakes. Again, no really unique values are attained by explosions, except when compared only with our screened population of apparently shallow events, in which case nearly all the explosions have higher values than any earthquake. Although spectral ratio shows promise of providing positive identification of explosions as such, we would rather not quote performance figures of this type based on our present data, since they would depend heavily on our single-station screening for depth based on pP. Earthquake identification percentages are given in Table II.

When we plot complexity against body-wave magnitude, we find a much better separation than when complexity is used alone. Figures 14 and 15 show such scatter

Conditions Defining Earthquake Population	All Data	Magnitude ≥ 4.8	Shallow Events	Shallow and Magnitude ≥ 4.8
Size of Earthquake Population	85	32	48	12
Earthquake Identification Percentages Using:				
Complexity Alone	36.5%	43.4%	22.9%	16.7%
Spectral Ratio Alone	57.6%	65.6%	62.5%	66.7%
Complexity vs Magnitude	84.7%	59.4%	87.5%	50.0%
Spectral Ratio vs Magnitude	95.2%	90.6%	100.0%	100.0%
Spectral Ratio vs Complexity vs Magnitude	98.8%	96.9%	100.0%	100.0%

TABLE II

EARTHQUAKE IDENTIFICATION PERCENTAGES

diagrams for two populations: all earthquakes and shallow earthquakes, respectively. The inadequacy of our data with respect to overlap in magnitude is clearly evident in Fig. 15. According to this plot, we could easily establish an earthquake-free region which also contains nearly all the explosions (type R_1 region), but it is clear that we would merely be discriminating on the basis of magnitude ! However, R_2 regions can be defined which seem not to be entirely functions of magnitude, and they are shown on Figs. 14 and 15 as the regions to the left of the dashed curves. The behavior of these curves below magnitude 4.9 is rather arbitrary, in view of our lack of explosion data, and we have drawn them in a way which is perhaps unjustifiably favorable to the performance of complexity as a discriminant. The performance figures in Table II are just the percentages of the various earthquake populations lying within these R_2 regions.

When spectral ratio is plotted versus magnitude, a real separation begins to appear. Figures 16 and 17 are analogous to 14 and 15 with respect to populations. Again we refrain from defining an R_1-type region and the associated explosion identification percentage, although the data is very encouraging that such a region exists. R_2-type regions are again outlined by dashed curves, and earthquake identification percentages are given in Table II. Note that the separation is perfect when our shallow population is used. Using the total population, four earthquakes cannot be distinguished from explosions. However, all four of these were reported by USCGS and assigned depths in excess of 50 km. Events of this kind would probably have been screened out

as deep by a network of stations on the basis of P-arrival times alone, and these four events in particular were judged to be deep from our single-station data on the basis of pP identification.

We have also plotted dominant period versus magnitude in order to assess the discrimination potential of this simple discriminant. Figure 18 is such a plot for the shallow earthquakes. A fairly good separation is shown here, but we are reluctant to claim performance figures because of the coarse quantization of the data (0.1 second increments) and the subjectivity of the measurement process itself. A very favorable interpretation of Fig. 18 would credit this display with correctly identifying 41 of the 48 earthquakes, i.e., 85.4%, while the corresponding percentage for the total earthquake population (not shown) is about 70%.

With our population of shallow events, there is no point in considering three dimensional discrimination, based on plotting spectral ratio versus complexity versus magnitude, since perfect separation is found in two dimensions. Since the separation is imperfect on the total population (because of the presence of deep events), it makes sense to try the three-dimensional plot in this case. The simplest way to use the three-dimensional data is to see if an event is identified as an earthquake by either the complexity versus magnitude or the spectral ratio versus magnitude criterion. Geometrically, this corresponds to extending both of the two-dimensional R_2 regions into the corresponding third dimension and then taking the union of the resulting three-dimensional regions. In general one could do better than this in three dimensions, but

with our data base it is unnecessary, since three of the four earthquakes which are ambiguous on the spectral ratio versus magnitude criterion are clearly rejected on the complexity versus magnitude criterion, and the fourth is hopelessly embedded in the explosion swarm of points. The reason these three events are rejected by complexity is that they all have a phase (picked as pP) in the second time window used for the complexity measurement, thus raising the value of that parameter. The fourth event is from Rumania, reported by USCGS at a depth of 158 km. Its pP phase is visible on our data at 38 seconds after P, just too late to influence the complexity. The three-dimensional discrimination percentages given in the first two columns of Table II represent identification of all but this one earthquake.

It is interesting to note that a set of five earthquakes, lying roughly on a line parallel to the dashed curve on our spectral ratio versus magnitude plot (Fig. 16) and very close to that line, are all rejected by the complexity versus magnitude criterion. All had pP phases on our data, and three were reported by USCGS at depths 26, 28 and 33 (R) km. One needs to be wary of the effect of PcP on complexity, but it appears that complexity of P waveform may be a useful adjunct to spectral ratio, even if it is only responding to pP.

We have not tried to derive a quantitative regional dependence of discrimination performance because the only well-defined region in our data is the Kurile-Kamchatka group of earthquakes and only Longshot represents an explosion in a similar kind of region. The non-Kurile-Kamchatka earthquakes are all continental in origin, but are

very widely distributed in location. However, we have seen from the histograms that these continental earthquakes have lower spectral ratios, as a group, and this is reflected in an even wider separation of the earthquake and explosion populations on a spectral ratio – magnitude plot. The elimination of Kurile-Kamchatka events has little effect on the performance of complexity, decreasing the 87.5% figure (for shallow events using a complexity – magnitude plot) to about 70%.

VI. CONCLUSIONS

We summarize our conclusions here in terms of the objectives outlined in the Introduction. We hesitate to draw any really firm conclusions from a study based upon such limited, unbalanced populations as is this one, but we feel that the following statements are justified by the facts.

A. General Effectiveness of the Short-Period Discriminants

Waveform complexity appears to be a discriminant of rather limited usefulness. Used alone it is probably not a satisfactory discriminant, in the sense discussed in Section IV. The results of the present study on this point are in remarkable agreement with earlier published results of ours,[8] on a different population, where an identification percentage of 15.8% was found. The complexity was defined differently in this earlier study, being an average value of the complexities measured on individual single-sensor waveforms from a number of subarrays. In that report, undue attention was perhaps drawn to the performance that would result (87%) if we ignored the Novaya Zemlya event. The analogous figure for our present population of shallow events is 62.5%. Since Novaya Zemlya is probably not the only source of complex explosions, it does not seem prudent to treat this event as an anomaly and forget it. Complexity looks much better when displayed as a function of magnitude, but our data is severely limited by the magnitude disparity and the scatter diagrams (Figs. 14 and 15) do not suggest that separation would be as good as the Table II values for a population including larger earthquakes and smaller explosions.

Spectral ratio, used alone, would not provide really useful separation by a single station, but when plotted against magnitude the results are most encouraging. The separation found here is complete, after screening for depth, and very nearly so for the total population. Moreover, the scatter diagrams (Figs. 16 and 17) suggest that the separation will persist with larger earthquakes and smaller explosions. Regional effects, if real, seem to work in our favor, by increasing the separation. We plan to pursue the study of this discriminant, extending the populations and varying the type of preprocessing used. We also plan to try variations and simplifications in the definition of spectral ratio.

B. Magnitude Thresholds for Short-Period Discriminants

In Section III we quoted the magnitude 4.3 as a threshold for reliable measurement of either of the short-period discriminants studied here. This is a strictly qualitative number, based on studying the waveforms and spectra of all the marginal events in our data base, and making a judgment on the degree to which the measurement seemed to be affected by the noise. Since the measurement quality is hardly determined by magnitude alone and is in any case very difficult to measure quantitatively, we feel that 4.5 is a reasonable threshold magnitude for discriminant measurement on Sino-Soviet earthquakes by LASA. It has been conservatively estimated[4] that surface waves are reliably detected at and above magnitude 4.9 for shallow Sino-Soviet earthquakes by LASA, hence we conclude that short-period discrimination by means of spectral ratio is potentially a very valuable supplement to the

established long-period technique. The average distance for earthquakes used in this study was 68°, while nearly all the explosions were at 83°. A station with the sensitivity of the Montana LASA at an average distance of 30° from its sources of events would have nearly a half-magnitude advantage in signal strength, and hence could push the measurement threshold down near 4.0. However, we do not know what to expect for the spectral ratios of explosions of that size, since this represents an extrapolation of our present data by a whole magnitude unit.

C. The Contribution of the Large Array to Short-Period Discrimination

In this study the large array was used only to form a preprocessed waveform of enhanced signal-to-noise ratio on which the discriminant measurements were made. The fact that we also used the resolving power of the array to provide epicenters is irrelevant to the evaluation of the array as an element in a world-wide network, since location (and to some extent depth) is determined by the network as a whole. Other ways of using the array in the discriminant-measuring process, such as measuring complexity and spectral ratio on subarray outputs and combining these, have not been considered here. Thus the discussion reduces to the characteristics of the single waveform produced by the array.

This waveform differs from the waveform produced by a single sensor or small array (such as a LASA subarray) in two ways. First, the signal-to-noise ratio is higher, by the equivalent of about three-quarters of a magnitude unit, on the array output due to the noise reduction resulting from the averaging of subarray waveforms.

Second, the signal component of the waveform is the average of the time-aligned signal waveforms from individual subarrays, and is therefore representative of that part of the signal waveform which is common to all these subarrays.

From the point of view of signal-to-noise ratio alone, the array can be thought of as a means of providing a single output trace of unusually low noise level. While preserving the signal level of a typical LASA single sensor, which is comparable to or slightly higher than other western U. S. stations (in one study,[6] LASA magnitudes averaged 0.2 units higher than USCGS magnitudes), the LASA beam output has a noise level of about[6] 0.2 to 0.3 mμ (r.m.s. value in the 0.6 to 2.0 Hz band). It is natural to compare the array to other techniques for obtaining high signal-to-noise ratios, such as the development of especially remote, quiet sites, deep holes or vertical arrays. It is important to compare the LASA to other such stations on the basis of signal-to-noise ratio, not signal or noise levels alone, after which it becomes a matter of how much signal-to-noise gain is obtained as a function of station cost, a comparison which is equally relevant to the detection problem, as well as the discrimination problem.

The second factor, however, relates specifically to the large geometrical aperture of the array. It is well known that event waveforms vary greatly from subarray to subarray within LASA, presumably since each waveform is influenced by different parts of the earth's crust and upper mantle under the array. We would like to think that the beam waveform is closer to that portion of the source waveform that

survives the long propagation path through the earth than is a single-sensor or subarray waveform. If this is so, then the beam waveform should be a more reliable one on which to measure discriminants than a waveform of comparable signal-to-noise ratio obtained from a single site. This conjecture can be tested only by carrying out the implied experiment, of which the present study represents one half. It would be very interesting to see the results of the other half, a large population study of short-period discriminant performance on single-station waveforms of comparable signal-to-noise ratio.

ACKNOWLEDGEMENT

An experimental study of this kind is inevitably the work of many people, in this case nearly the entire Lincoln Laboratory Seismic Discrimination Group. Most of the original planning and early-phase organization was done by H. W. Briscoe, R. M. Sheppard and R. T. Lacoss were responsible for the individual measurements, aided by N. E. Williams and R. Walsh. Analysis of the resulting data base was done by the author of the report.

The recordings forming the basis of this work were also used in a study of long-period discrimination by J. Capon and R. J. Greenfield, and their results are included in the report Reference 4.

REFERENCES

1. J. N. Brune, A. Espinosa and J. Oliver, "Relative Excitation of Surface Waves by Earthquakes and Underground Explosions in the California-Nevada Region," J. Geophys. Res., 68, June 1963.

2. R. C. Liebermann, et al, "Excitation of Surface Waves by the Underground Nuclear Explosion Longshot," J. Geophys. Res., 71, September 1966.

3. P. E. Green, R. A. Frosch and C. F. Romney, "Principles of an Experimental Large Aperture Seismic Array (LASA)," Proc. IEEE, 53, December 1965.

4. J. Capon, R. J. Greenfield and R. T. Lacoss, "Long-Period Signal Processing Results for Large Aperture Seismic Array," Lincoln Laboratory, M. I. T., Technical Note 1967-50, 15 November 1967.

5. Seismic Discrimination, Semiannual Technical Summary Report, Lincoln Laboratory, M. I. T., 30 June 1967, p. 5.

6. E. J. Kelly, "LASA On-Line Detection, Location and Signal-to-Noise Enhancement," Lincoln Laboratory, M. I. T., Technical Note 1966-36, 1 July 1966.

7. E. W. Carpenter, "Teleseismic Methods for the Detection, Identification and Location of Underground Explosions," VESIAC Report 4410-67-X, University of Michigan, April 1964.

8. Seismic Discrimination, Semiannual Technical Summary Report, Lincoln Laboratory, M. I. T., 31 December 1966, p. 7.

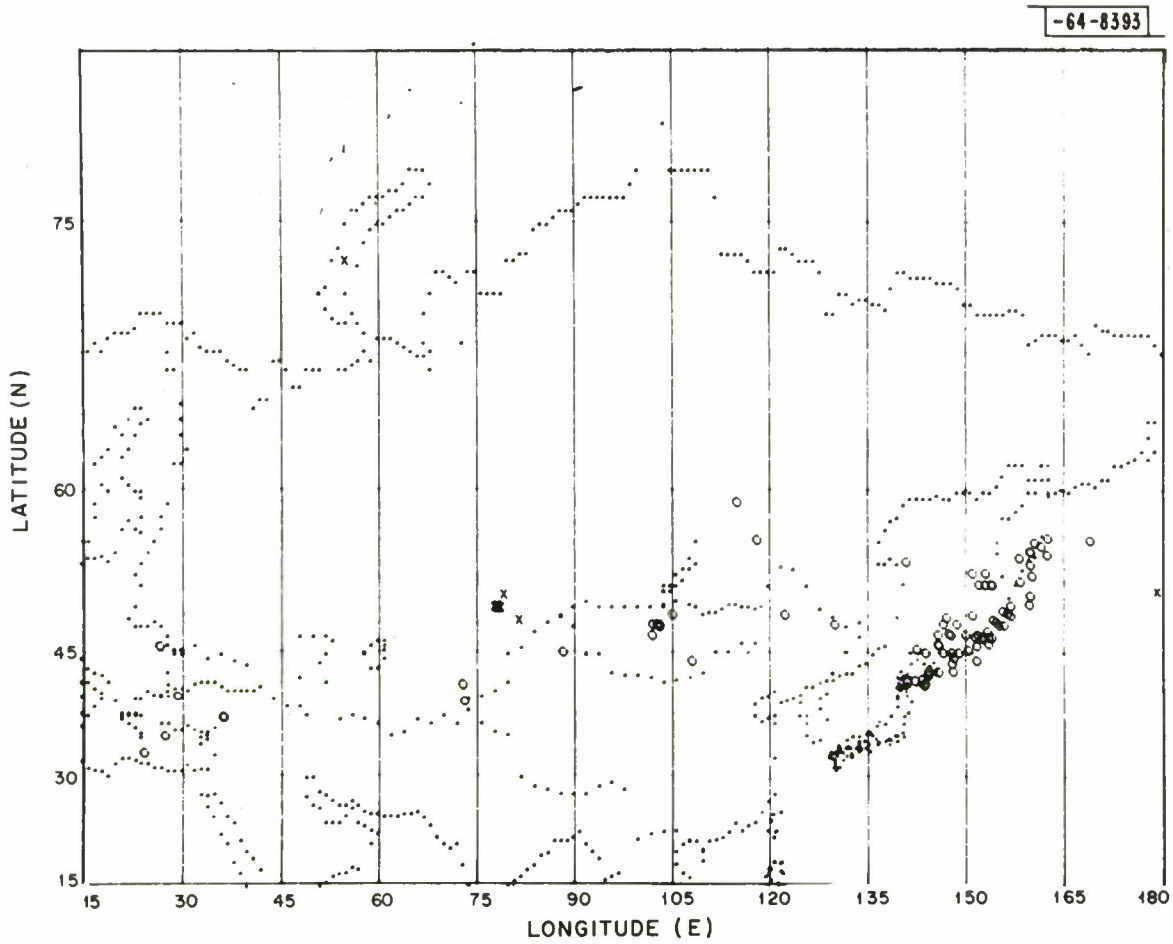

Fig. 1. Geographical distribution of earthquakes (O) and explosions (X).

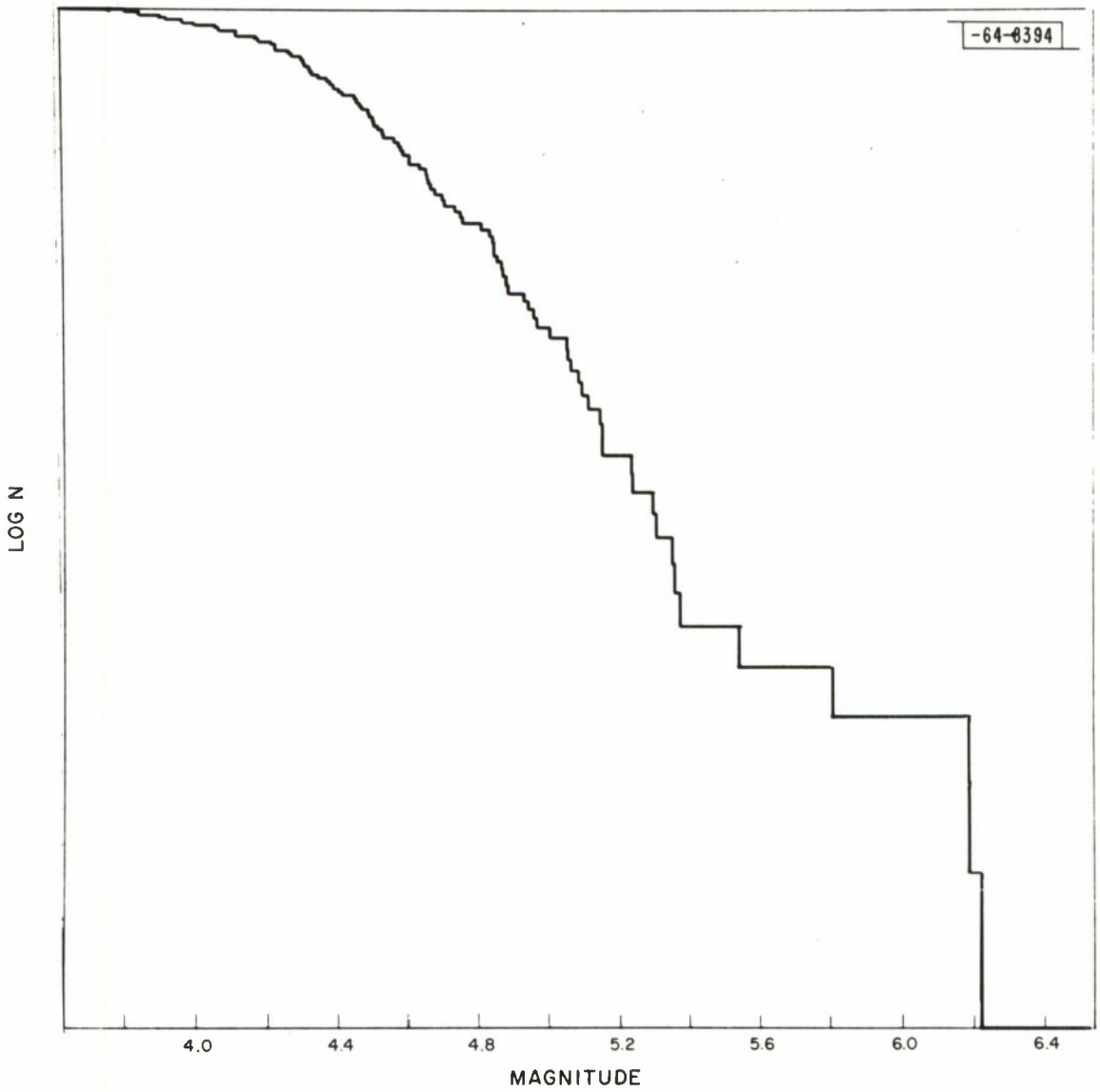

Fig. 2. Cumulative seismicity curve for the earthquake population. (Logarithm of the number of earthquakes exceeding a given magnitude vs magnitude.)

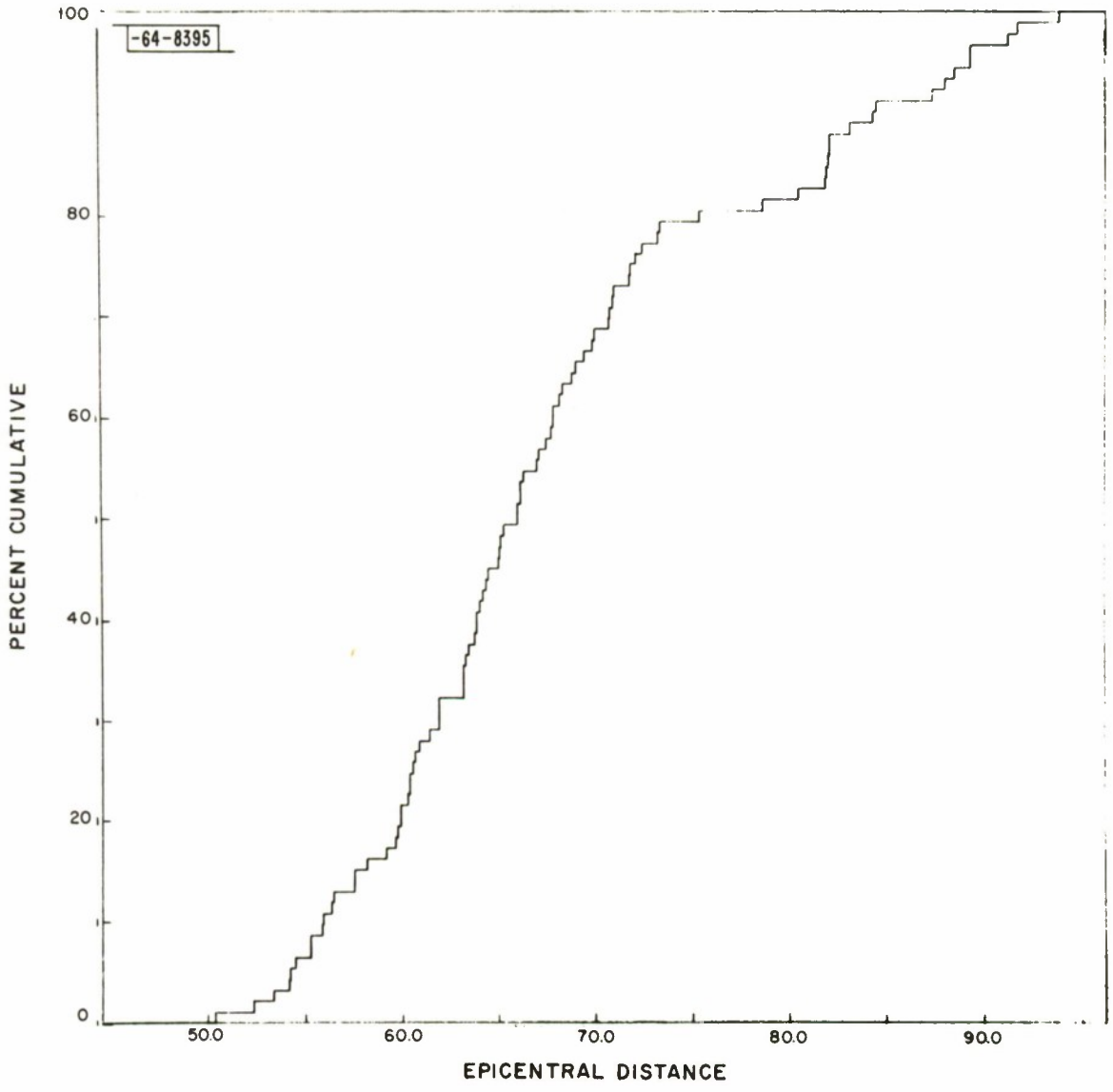

Fig. 3. Cumulative histogram of epicentral distance for earthquakes.

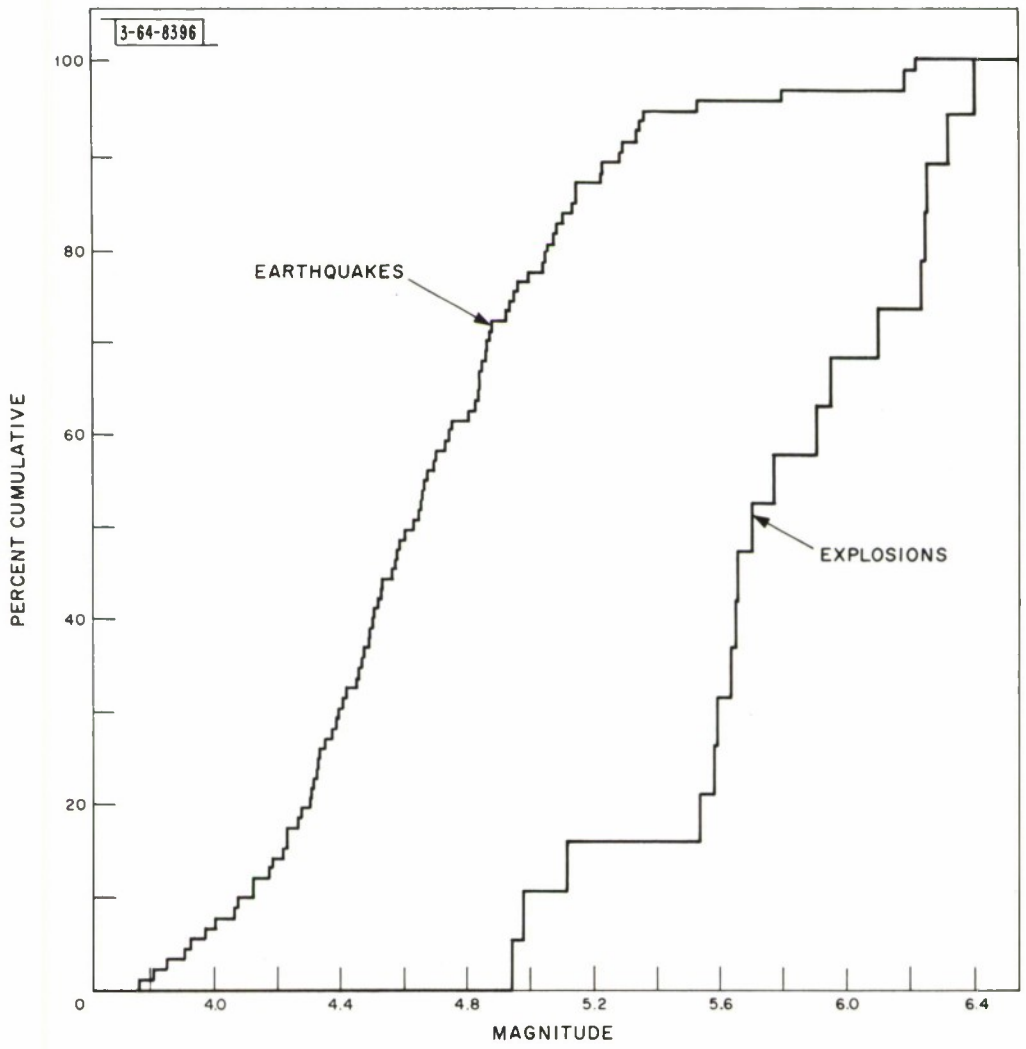

Fig. 4. Cumulative histograms of measured magnitude for earthquakes and explosions.

-64-8397

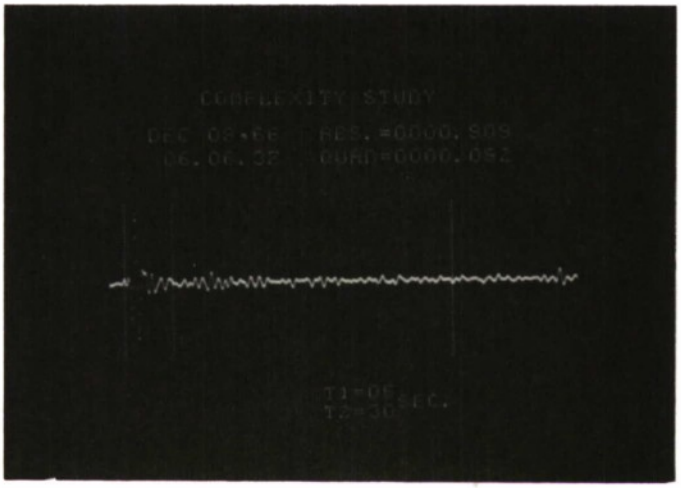

Fig. 5. Example of a waveform illustrating the complexity measurement.

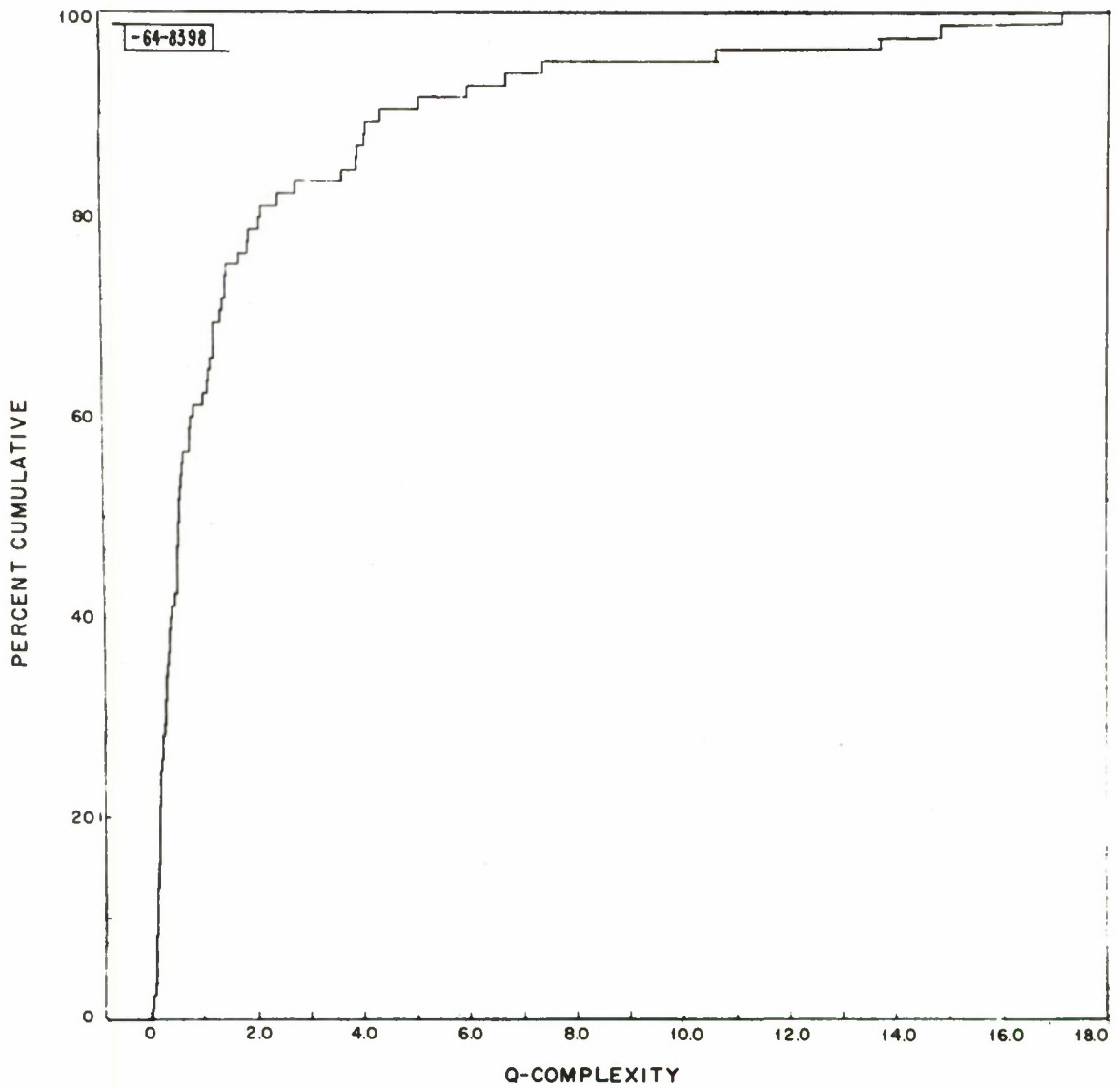

Fig. 6. Cumulative histogram of quadratic complexity for earthquakes.

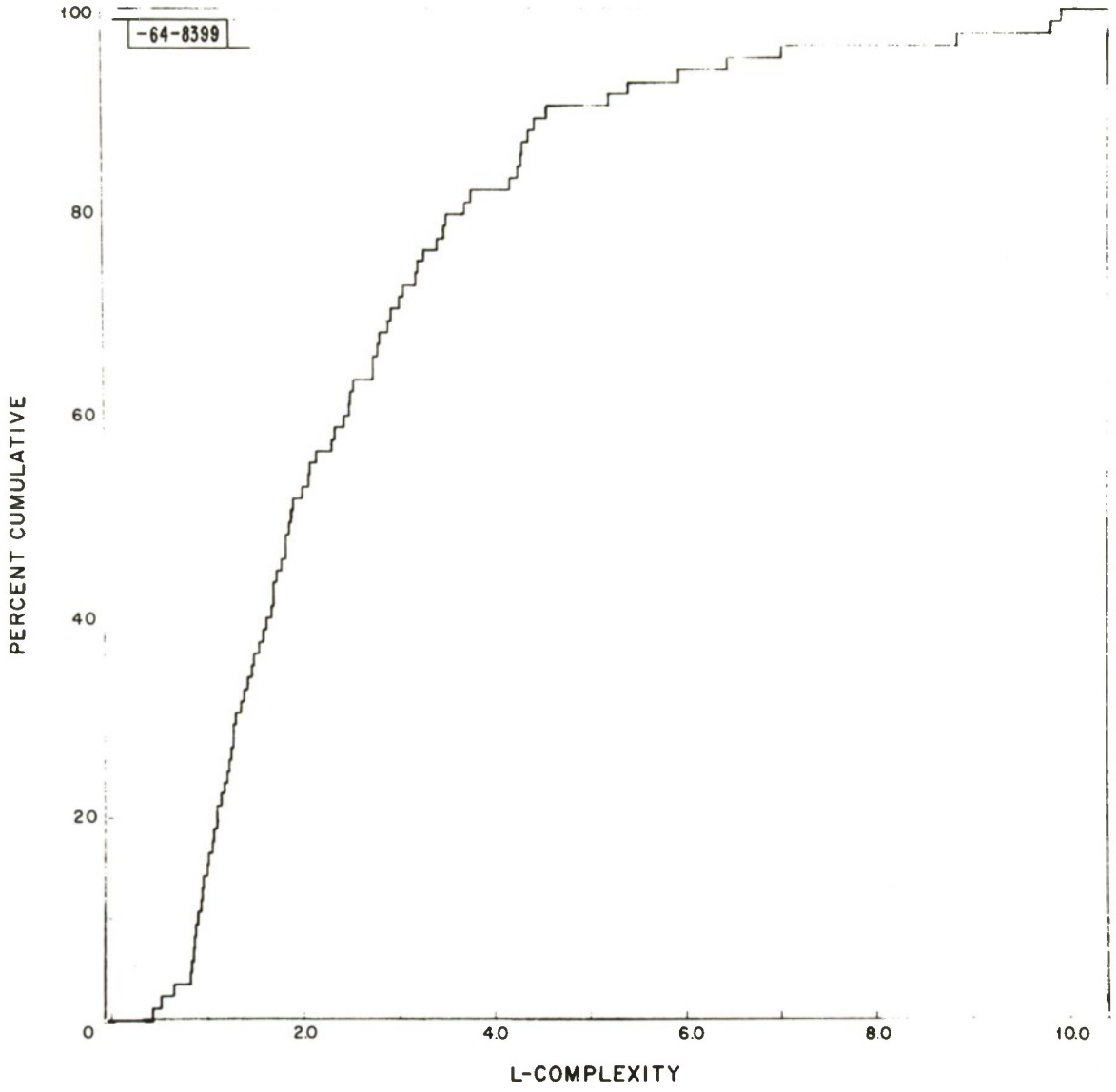

Fig. 7. Cumulative histogram of linear complexity for earthquakes.

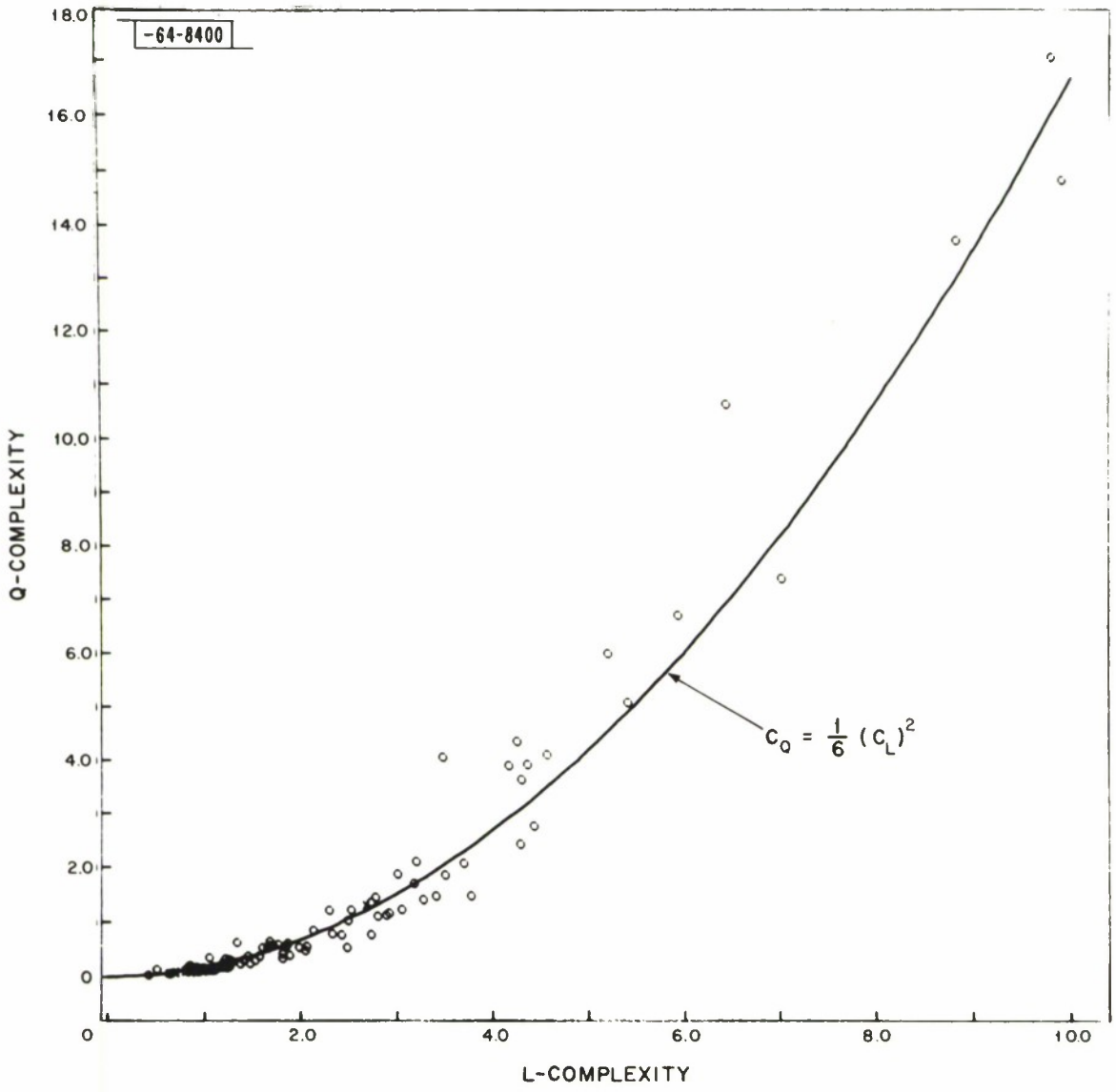

Fig. 8. Quadratic complexity vs linear complexity for earthquakes (O) and explosions (X).

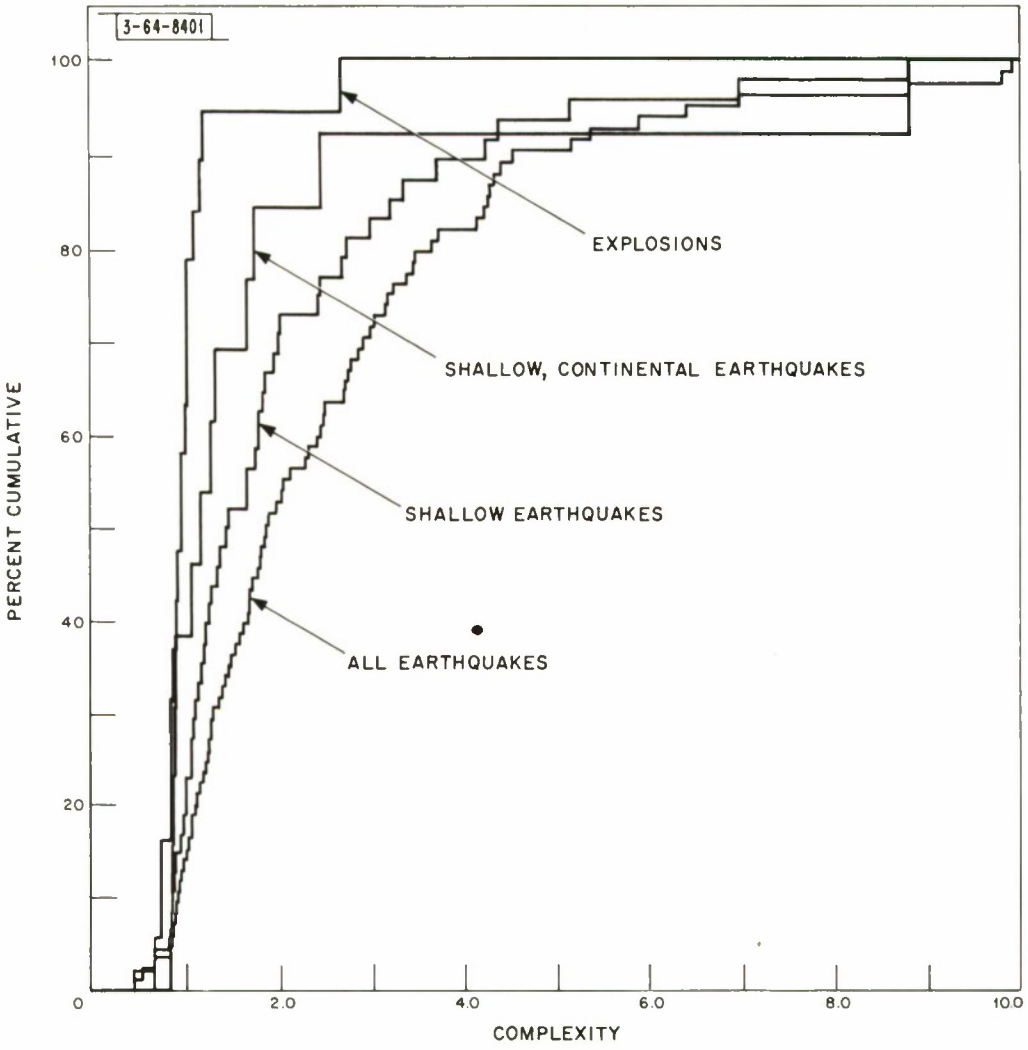

Fig. 9. Cumulative histograms of complexity for explosions and three earthquake populations.

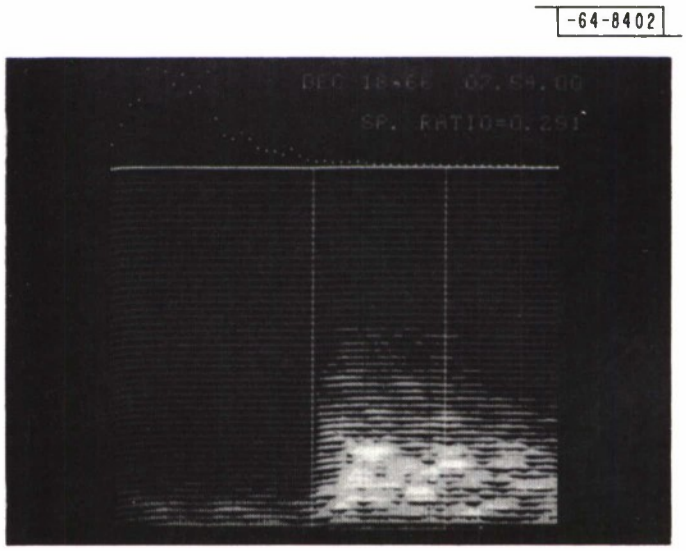

Fig. 10. Example of a sonogram illustrating the spectral ratio measurement.

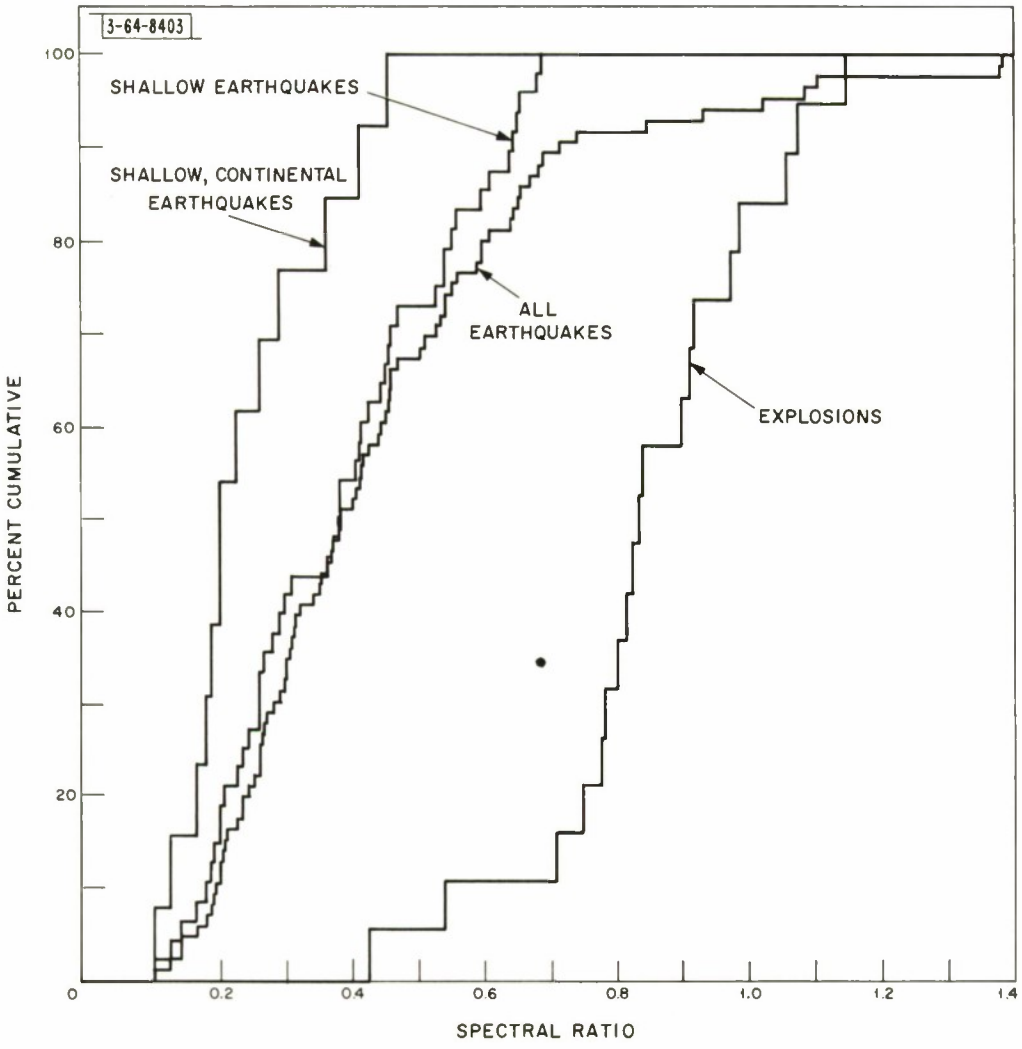

Fig. 11. Cumulative histograms of spectral ratio for explosions and three earthquake populations.

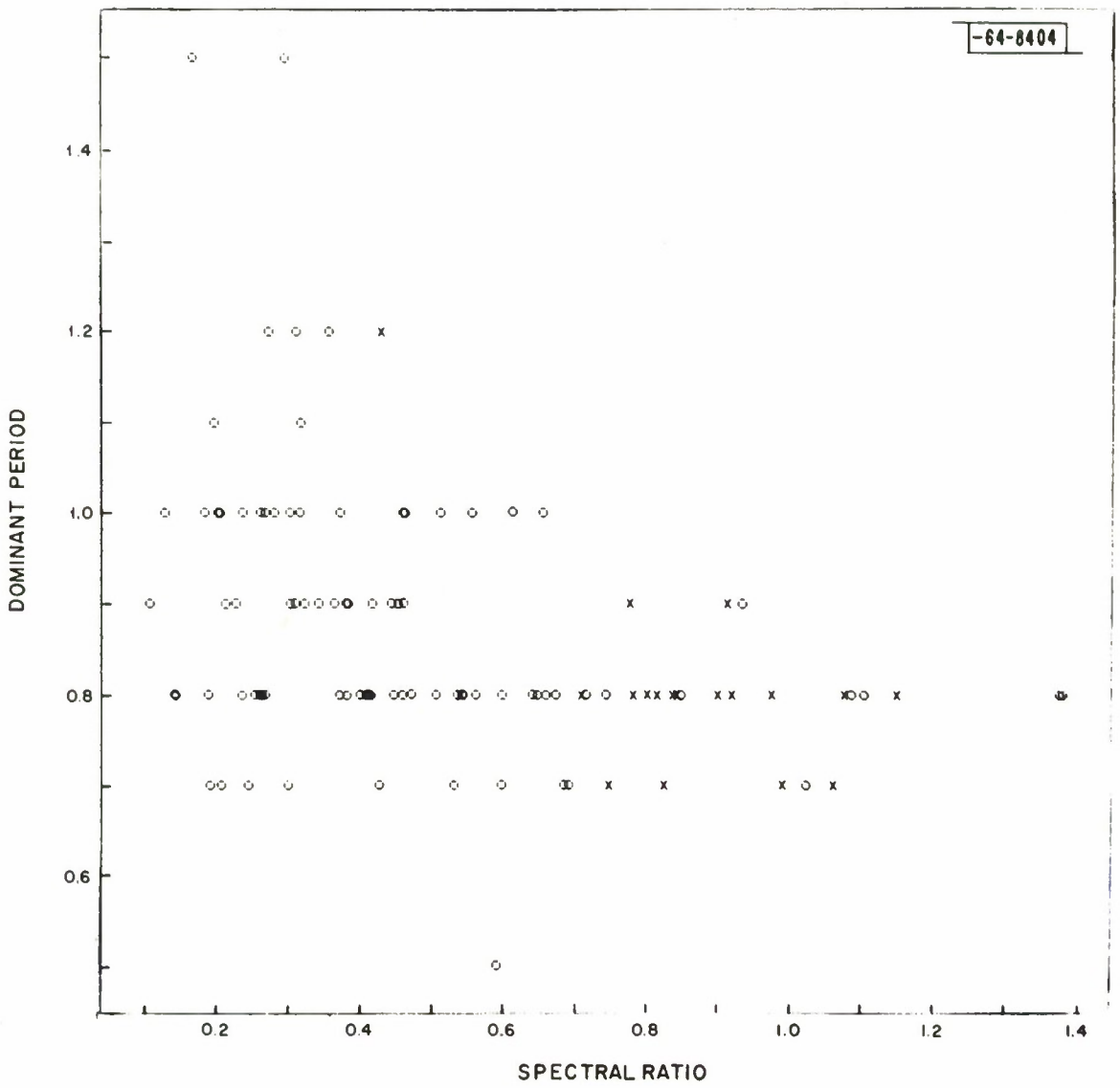

Fig. 12. Dominant period vs spectral ratio for 85 earthquakes (O) and 19 explosions (X).

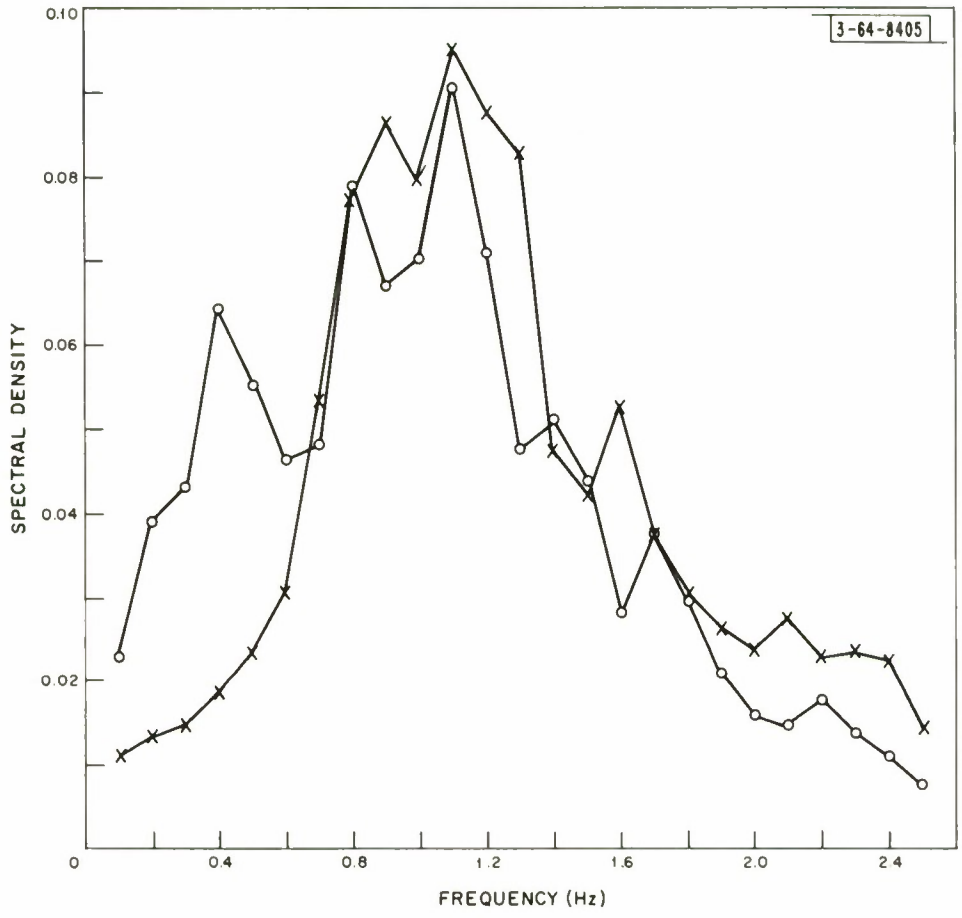

Fig. 13. Typical spectral densities for an earthquake (O) and an explosion (X). (Arbitrary vertical scale.)

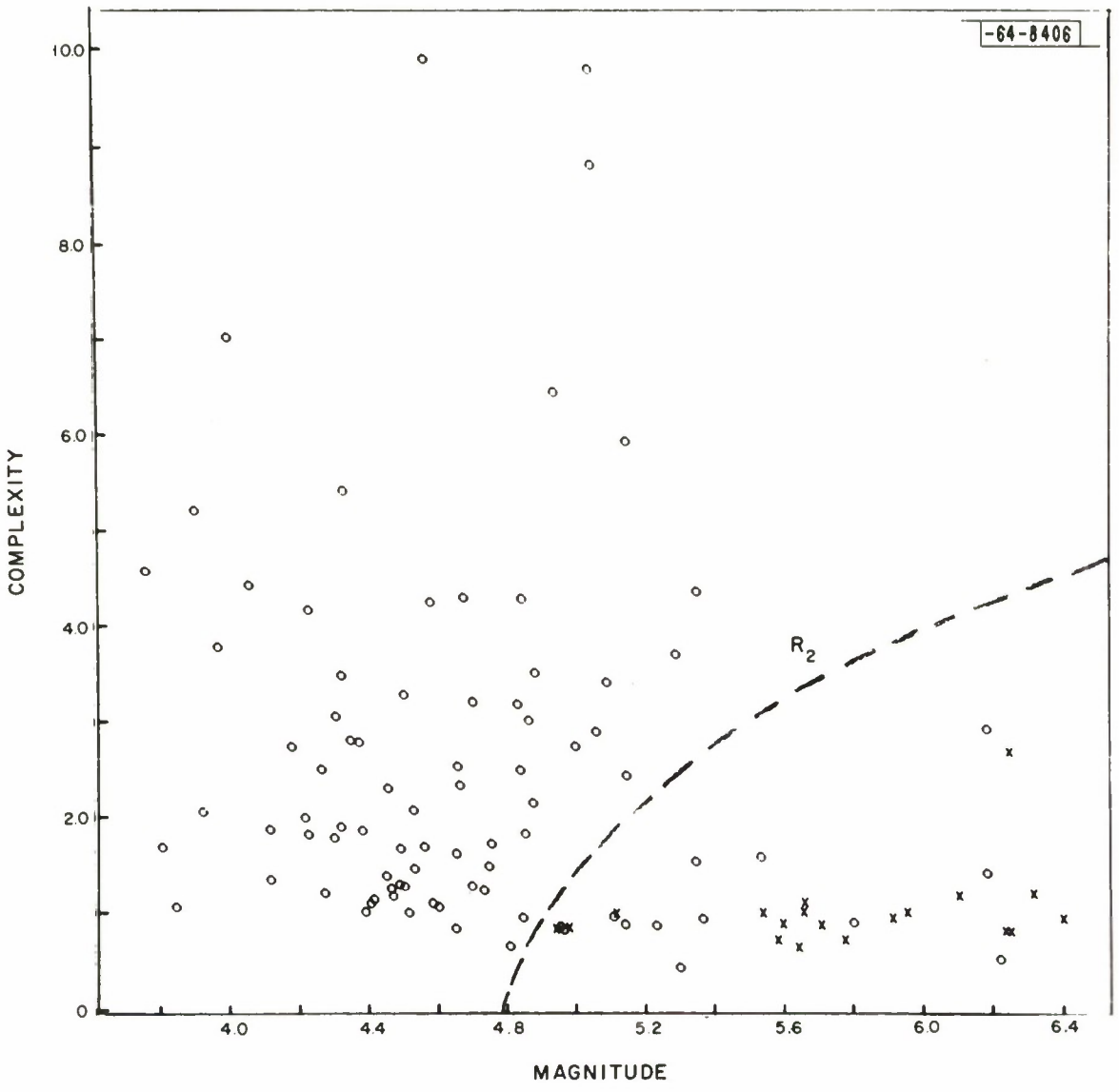

Fig. 14. Complexity vs magnitude for 85 earthquakes (O) and 19 explosions (X).

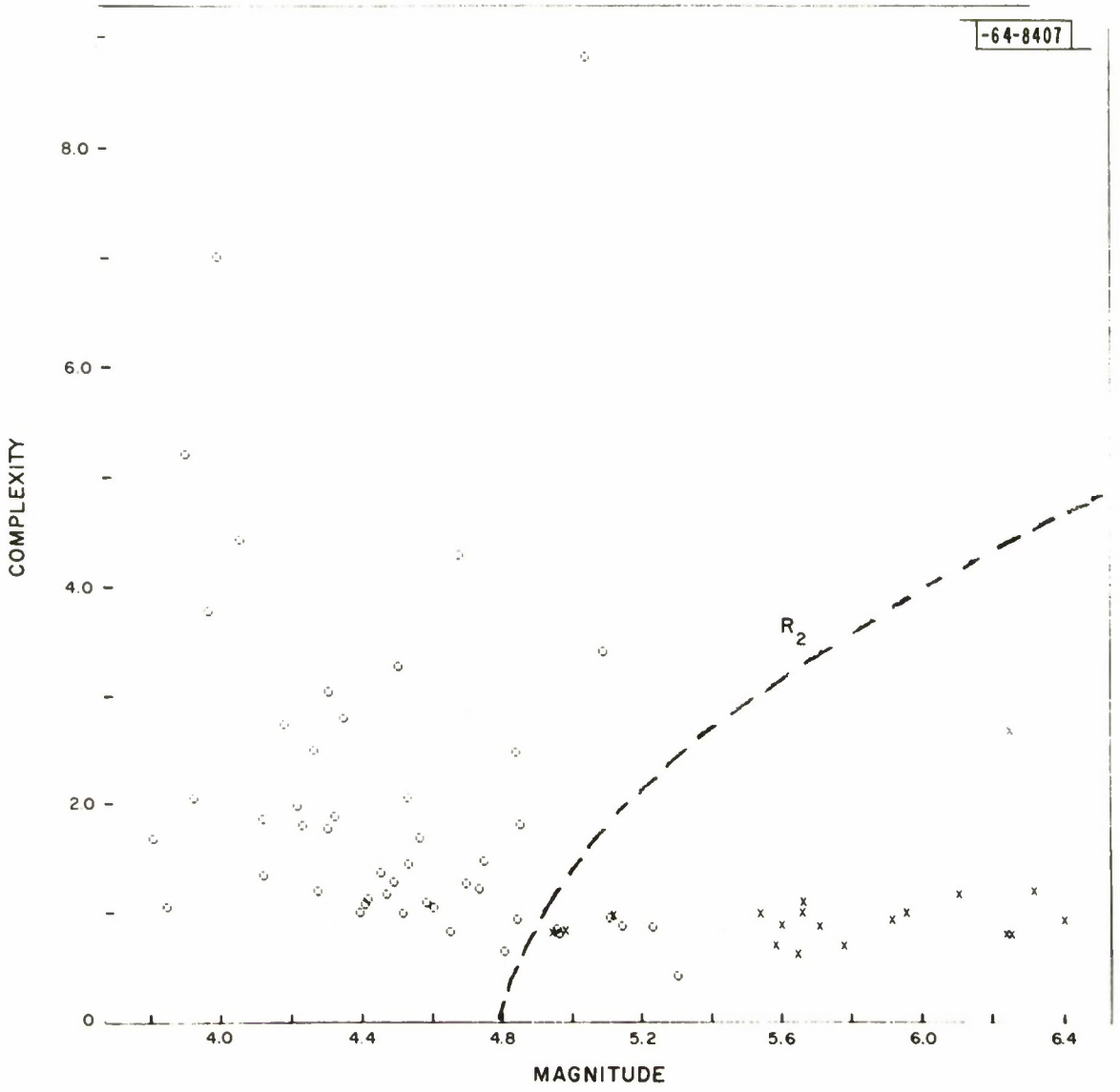

Fig. 15. Complexity vs magnitude for 48 shallow earthquakes (O) and 19 explosions (X).

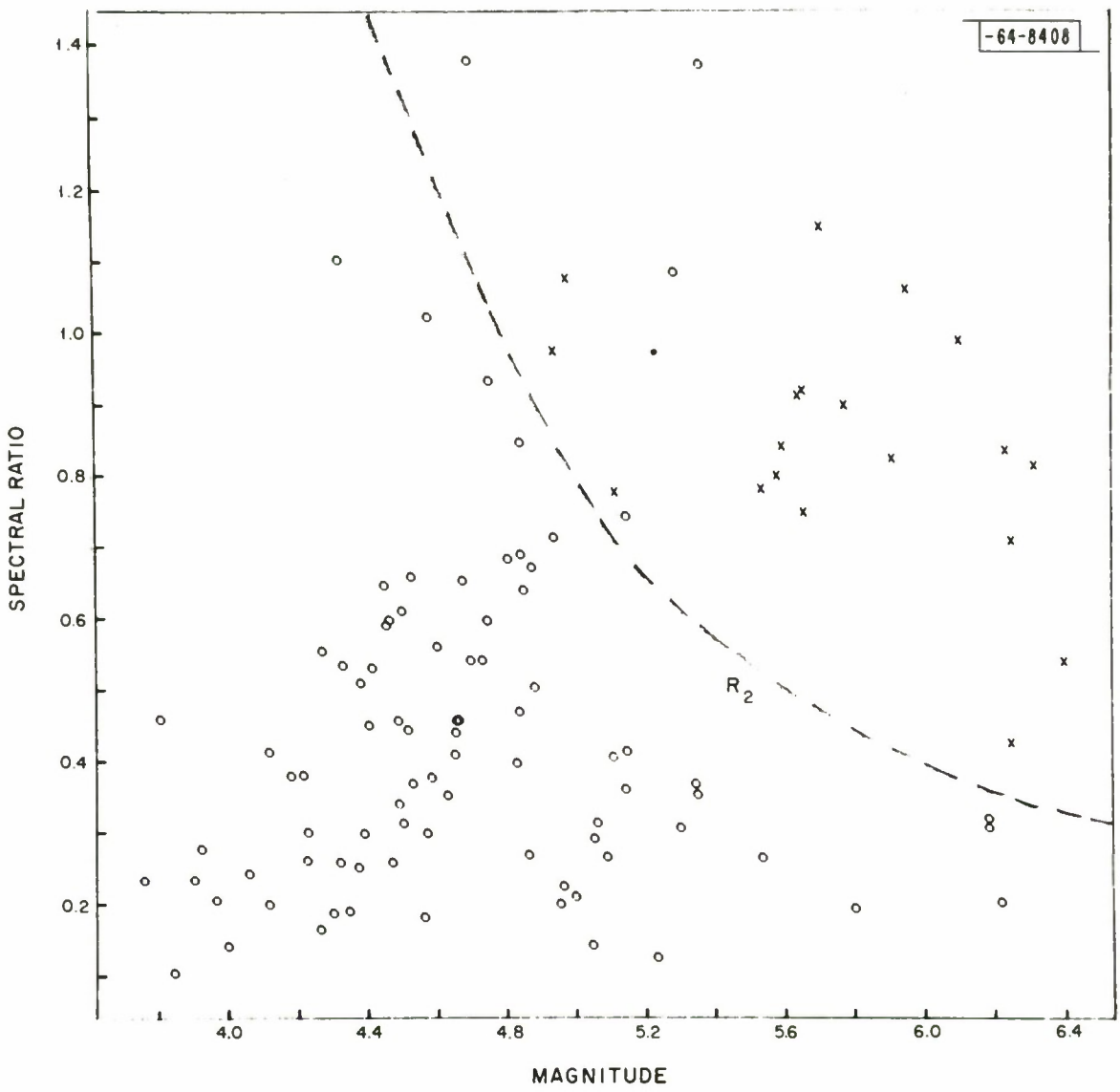

Fig. 16. Spectral ratio vs magnitude for 85 earthquakes (O) and 19 explosions (X).

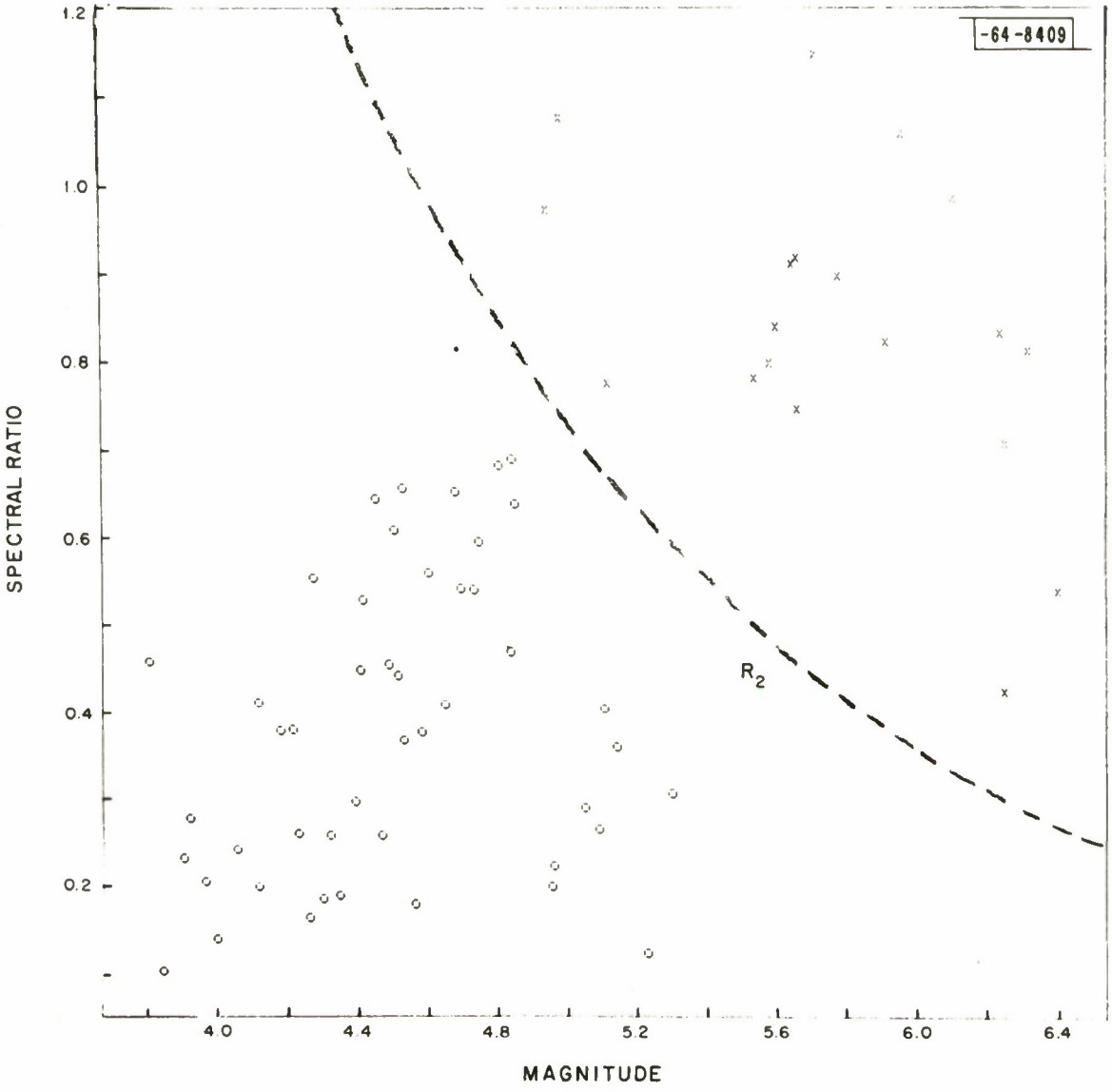

Fig. 17. Spectral ratio vs magnitude for 48 shallow earthquakes (O) and 19 explosions (X).

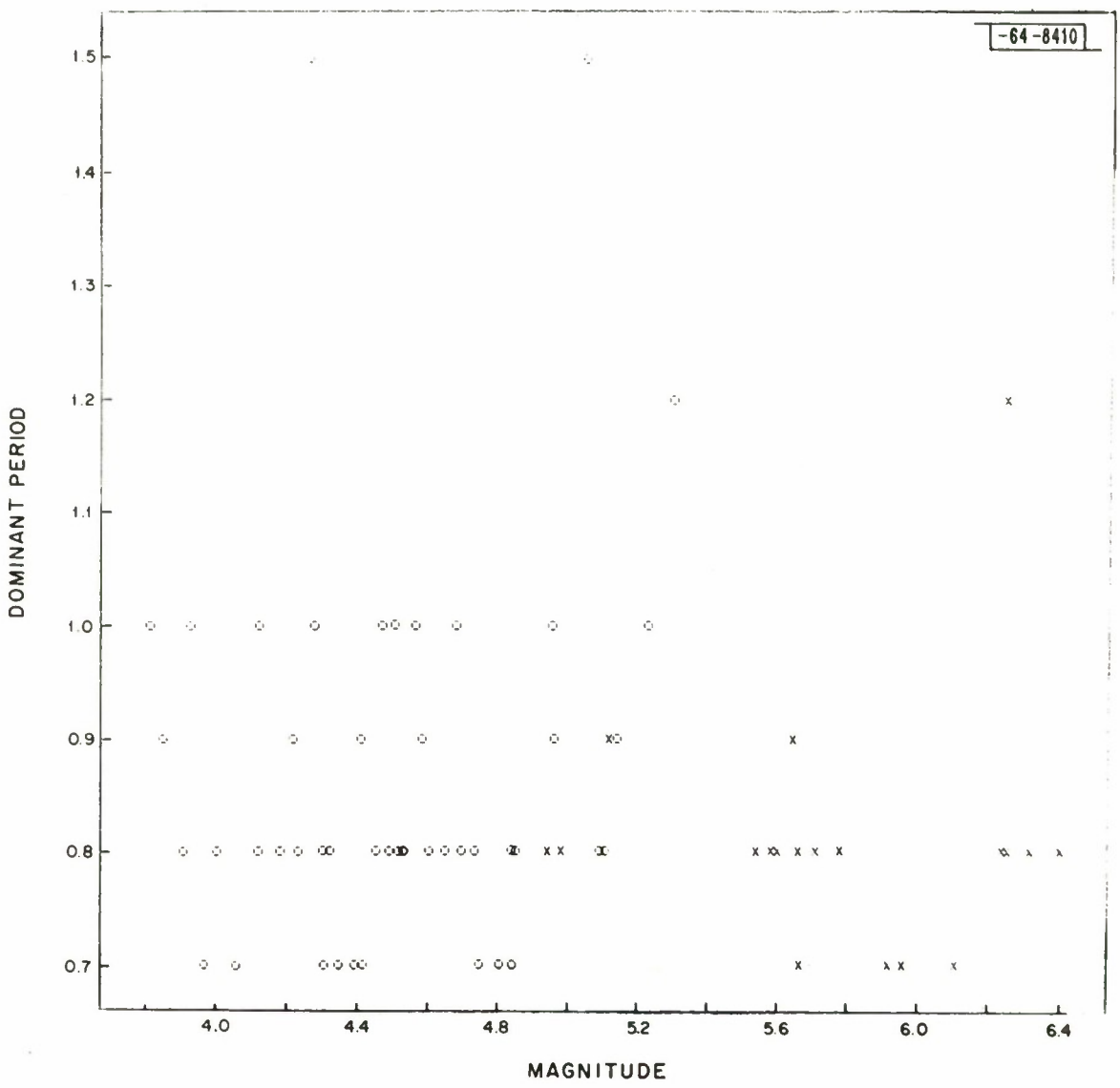

Fig. 18. Dominant period vs magnitude for 48 shallow earthquakes (O) and 19 explosions (X).

APPENDIX

The following is a list of the earthquakes used in this study. If a PDE number is given, then all data is taken from the PDE card (card 80/43 refers to the forty-third event on card 80 of the year in question). Otherwise the data is obtained from the LASA recording, as described in the text.

Date	Origin (GMT)	Lat. (N)	Long. (E)	Depth (Km.)	Mag.	PDE
1966						
12 Nov	12 49 43.6	41.7	144.1	33 (R)	5.8	80/43
12 Nov	13 28 23*	41.5	143.8	54	4.1	88/2
16 Nov	14 57 18.1	56.0	169.0	106	4.3	
17 Nov	19 27 05*	46.1	153.6	33 (R)	4.4	81/34
18 Nov	08 46 11.7	50.0	157.0	33	4.0	
19 Nov	02 02 52.9	48.1	148.7	33	4.6	
19 Nov	13 19 30.6	49.2	156.5	33	4.3	
21 Nov	12 19 27.3	46.7	152.5	40 (R)	5.6	88/6
22 Nov	06 29 53.5	48.1	146.7	453 (R)	5.6	82/40
24 Nov	02 56 36.6	42.9	146.0	33	3.9	
25 Nov	19 32 34.9	45.9	146.0	33	4.8	
25 Nov	20 31 36*	41.5	72.6	33 (R)	5.1	88/10
26 Nov	17 05 08*	42.6	144.5	54	3.9	89/10
27 Nov	00 20 18.6	52.0	153.0	33	4.6	
27 Nov	11 01 10*	49.5	155.8	40	4.9	85/41
27 Nov	12 48 02*	48.1	155.0	28	4.5	84/42

27 Nov	18 00 55.8	45.0	144.0	33	4.7	
29 Nov	08 09 39.9	55.0	160.0	33	4.4	
29 Nov	14 08 13.8	42.1	143.4	33	4.1	86/54
30 Nov	00 04 36*	46.9	152.7	33 (R)	4.4	88/11
30 Nov	08 49 57.5	46.8	154.1	33	4.1	
1 Dec	11 38 43.8	47.1	147.5	33	4.2	
2 Dec	01 39 49.2	54.4	158.2	33	3.9	
3 Dec	05 21 05.9	43.8	148.0	33	4.2	
4 Dec	11 06 17.7	55.5	161.7	33	4.3	
5 Dec	05 13 31.4	45.4	142.6	33	4.4	
6 Dec	07 18 40*	50.1	159.8	27	5.4	85/56
6 Dec	10 45 02.0	41.8	141.1	127	4.3	85/57
7 Dec	04 15 22*	46.8	153.6	49	4.5	92/17
7 Dec	15 12 44.5	49.0	157.0	33	4.7	
7 Dec	17 17 42.0	44.2	151.7	26	5.8	85/59
8 Dec	05 51 18.1	45.0	148.0	33	4.1	
8 Dec	13 02 46.0	43.0	148.2	33	4.0	
9 Dec	05 56 33.0	52.0	154.0	33	4.8	
11 Dec	19 47 34.2	42.9	144.5	57	4.8	87/43
12 Dec	10 35 57.1	47.0	151.8	33	4.5	
14 Dec	14 49 59.8	45.6	26.3	158	4.8	87/46
18 Dec	07 42 18.8	35.1	27.1	33 (R)	4.7	90/32
21 Dec	02 30 11.4	47.0	148.0	33	4.5	
21 Dec	11 41 50.1	33.0	24.0	33	4.3	
22 Dec	12 10 06*	52.3	158.5	61	4.4	93/14
22 Dec	17 26 32*	48.8	147.2	38	4.5	93/15
22 Dec	19 24 06.5	48.6	154.3	77 (R)	5.2	90/41
23 Dec	14 05 54.3	47.0	150.0	33	4.7	

23 Dec	23 37 39.5	53.8	160.0	26	4.4	
23 Dec	23 49 27*	54.7	162.5	28	4.9	91/33
30 Dec	01 57 09.7	40.0	29.0	33	4.3	
30 Dec	04 40 07*	52.8	160.3	33 (R)	4.8	96/24
31 Dec	00 29 14.0	48.0	130.0	31	4.5	
31 Dec	00 30 08.6	56.0	118.0	33	4.6	
31 Dec	04 53 56.6	49.0	122.4	33	4.5	
31 Dec	06 51 25.5	51.0	160.0	33	4.3	
1967						
5 Jan	02 10 00.4	49.0	105.0	33	3.8	
5 Jan	04 54 22.4	44.0	108.0	13	4.5	
5 Jan	10 07 58.3	39.4	72.9	11	5.3	1/11
7 Jan	10 49 16.7	59.0	115.0	33	4.7	
8 Jan	18 31 59.7	56.2	162.7	24	4.9	3/18
11 Jan	06 56 42.4	46.9	101.8	33	4.1	
18 Jan	06 21 27.0	45.3	150.6	33 (R)	4.3	7/31
18 Jan	08 29 03.4	42.0	142.4	65	4.8	4/31
18 Jan	11 17 44.7	48.0	156.0	33	4.8	
20 Jan	03 27 13.9	48.0	103.0	33 (R)	5.0	7/36
20 Jan	06 23 16.3	47.8	103.1	33 (R)	5.0	6/17
22 Jan	12 01 49.0	48.0	102.1	33 (R)	5.1	6/21
22 Jan	12 16 02*	48.0	102.8	33 (R)	5.0	7/43
27 Jan	06 22 29.5	52.0	154.0	65	4.5	
1 Feb	09 18 50.5	55.8	160.7	140 (R)	4.4	10/47
22 Feb	14 50 33.1	48.3	154.7	45	4.7	12/57
24 Feb	04 32 21.6	53.0	151.0	20	4.5	
5 Mar	09 55 15*	46.8	152.7	33 (R)	4.4	16/36

10 Mar	09 02 20.5	45.0	146.7	33	4.5	
19 Mar	03 36 49.8	45.0	149.0	47	4.2	
29 Mar	10 01 10*	44.4	148.4	26	4.4	26/22
30 Mar	15 32 24.8	45.9	146.0	33	5.1	
1 Apr	14 00 33.8	45.8	151.7	23	5.4	21/60
5 Apr	08 08 49.5	52.0	152.0	33	4.5	
7 Apr	17 07 16.2	37.4	36.1	49	4.8	27/23
7 Apr	17 39 47.7	37.4	36.1	33	4.5	
7 Apr	19 39 13*	47.0	146.0	296	5.0	28/22
8 Apr	08 55 40*	47.3	153.4	60 (R)	4.6	31/6
22 Apr	05 18 53.8	45.0	88.0	17	4.6	
22 Apr	23 00 32*	46.8	151.6	73	4.4	31/26
26 Apr	16 47 06.1	53.0	153.0	33	4.4	
28 Apr	03 07 20.2	49.0	151.0	33	3.8	
28 Apr	08 56 41.1	54.0	141.0	33	4.4	

DOCUMENT CONTROL DATA – R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Lincoln Laboratory, M. I. T.	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE

A Study of Two Short-Period Discriminants

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Note

5. AUTHOR(S) *(Last name, first name, initial)*

Kelly, Edward J.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
12 February 1968	60	8

8a. CONTRACT OR GRANT NO. AF 19(628)-5167 b. PROJECT NO. ARPA Order 512 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Note 1968-8 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* ESD-TR-68-16

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Department of Defense

13. ABSTRACT

This paper is the result of a large-population study of discrimination between earthquakes and explosions, using only short-period data from a single station. The data was obtained from the Large Aperture Seismic Array and the two discriminants studied were waveform complexity and spectral ratio. Procedures for multivariate discrimination are developed and results are given in terms of earthquake identification percentages using these discriminants singly, and in combination with each other and with magnitude. The results are quite encouraging, complete separation being found for explosions and shallow earthquakes using spectral ratio and magnitude together.

14. KEY WORDS

seismic array
seismology
earthquakes
earth sensors
large-aperture seismic array
Montana